Solution of assignment 12, ST2304

Problem 1

To calculate the sample size needed to detect a difference between the probability of a binary outcome in two groups we use the in build function `power.prop.test()`. In the help file we see

Usage:

```
power.prop.test(n = NULL, p1 = NULL, p2 = NULL, sig.level = 0.05,
                power = NULL,
                alternative = c("two.sided", "one.sided"),
                strict = FALSE)
```

Arguments:

```
     n: Number of observations (per group)

    p1: probability in one group

    p2: probability in other group

sig.level: Significance level (Type I error probability)

  power: Power of test (1 minus Type II error probability)

alternative: One- or two-sided test

  strict: Use strict interpretation in two-sided case
```

Details:

```
  Exactly one of the parameters 'n', 'p1', 'p2', 'power', and
  'sig.level' must be passed as NULL, and that parameter is
  determined from the others. Notice that 'sig.level' has a non-NULL
  default so NULL must be explicitly passed if you want it computed.
```

Thus, in this case, since we want to calculate n, we have to specify p_1, p_1 and `power`, and we can use the default setting for α `sig.level=0.05`

```
> power.prop.test(p1=0.001, p2=0.002,power=0.9)

    Two-sample comparison of proportions power calculation

              n = 31472.91   <---
             p1 = 0.001
             p2 = 0.002
      sig.level = 0.05
          power = 0.9
    alternative = two.sided

 NOTE: n is number in *each* group
```

Thus, when $p_1 = 0.001$ and $p_2 = 0.002$ we need a sample size of at least 31473 in each group.

Similarly, when $p_1 = 0.01$ and $p_2 = 0.02$, the sample size in each group needs to be at least 3103, and 266 when $p_1 = 0.01$ and $p_2 = 0.02$.

So even though in all cases the probability in the second group is twice the probability in the first group, we need a larger sample size when the probability is smaller. In order to detect a rare event, you need a large number of observations, and to estimate the frequency of that event, you need a still larger number. For example in the first case, you expect to observe 31473*0.001 = 31 occurences of the event in the first group and 62 in the second group, which seems to be the minimum to determine whether or not the probabilities are different.

Problem 2

1. If we look at the help of `power.t.test` we see

   Usage:

   ```
   power.t.test(n = NULL, delta = NULL, sd = 1, sig.level = 0.05,
                power = NULL,
                type = c("two.sample", "one.sample", "paired"),
                alternative = c("two.sided", "one.sided"),
                strict = FALSE)
   ```

   Arguments:

   ```
         n: Number of observations (per group)

     delta: True difference in means

        sd: Standard deviation

   sig.level: Significance level (Type I error probability)

     power: Power of test (1 minus Type II error probability)

      type: Type of t test

   alternative: One- or two-sided test

    strict: Use strict interpretation in two-sided case
   ```

   Details:

   ```
       Exactly one of the parameters 'n', 'delta', 'power', 'sd', and
       'sig.level' must be passed as NULL, and that parameter is
       determined from the others. Notice that the last two have non-NULL
       defaults so NULL must be explicitly passed if you want to compute
       them.
   ```

   This means that if we want to determine the power, we need to provide arguments to all of `n`, `delta`, `sd` and `sig.level`, where `delta` is the true difference in means $\mu_1 - \mu_2$. For example, for $\mu_1 - \mu_2 = 0.5$ we get

```
> power.t.test(n=20,delta=0.5,sd=1,type="two.sample",alternative="one.sided")

     Two-sample t test power calculation

              n = 20
          delta = 0.5
             sd = 1
      sig.level = 0.05
          power = 0.4633743 <---
    alternative = one.sided

 NOTE: n is number in *each* group
```

In this case, the power is 0.463. For plotting a curve, we only need this single number. In the help section, we see that

```
Value:

    Object of class "power.htest", a list of the arguments
    (including the computed one) augmented with 'method' and 'note'
    elements.
```

The relevant point here is that the output of `power.t.test()` is a list. Each item in a list has a name, and the item can be retrieved from the list using either `list$name` or `list[["name"]]`. In this case,

```
> power.t.test(n=20,delta=0.5,sd=1,type="two.sample",
alternative="one.sided")$power
[1] 0.4633743
> power.t.test(n=20,delta=0.5,sd=1,type="two.sample",
alternative="one.sided")[["power"]]
[1] 0.4633743
```

Recall that `curve(expr(x),from=a, to=b)` draws a curve by varying x between a and b and evaluating the expression for those values. In this case, we want to vary the difference between the means $\mu_1 - \mu_2$, so we set `delta=x`

```
curve(power.t.test(n=20,delta=x,sd=1,type="two.sample",
alternative="one.sided")$power,
from=-2, to=4, xlab="difference between the means", ylab="power")
```

Instead of the inbuild function, we could have made our own function, similar to the one in section 4.5 of handout 5. However, this takes much longer to compute each value - several seconds versus a fraction of each second.

```
Power.t.test <- function(mu1, mu2, m=1000) {
  n.rejections <- 0
  for (i in 1:m) {
    X1 <- rnorm(20,mu1,1)
```
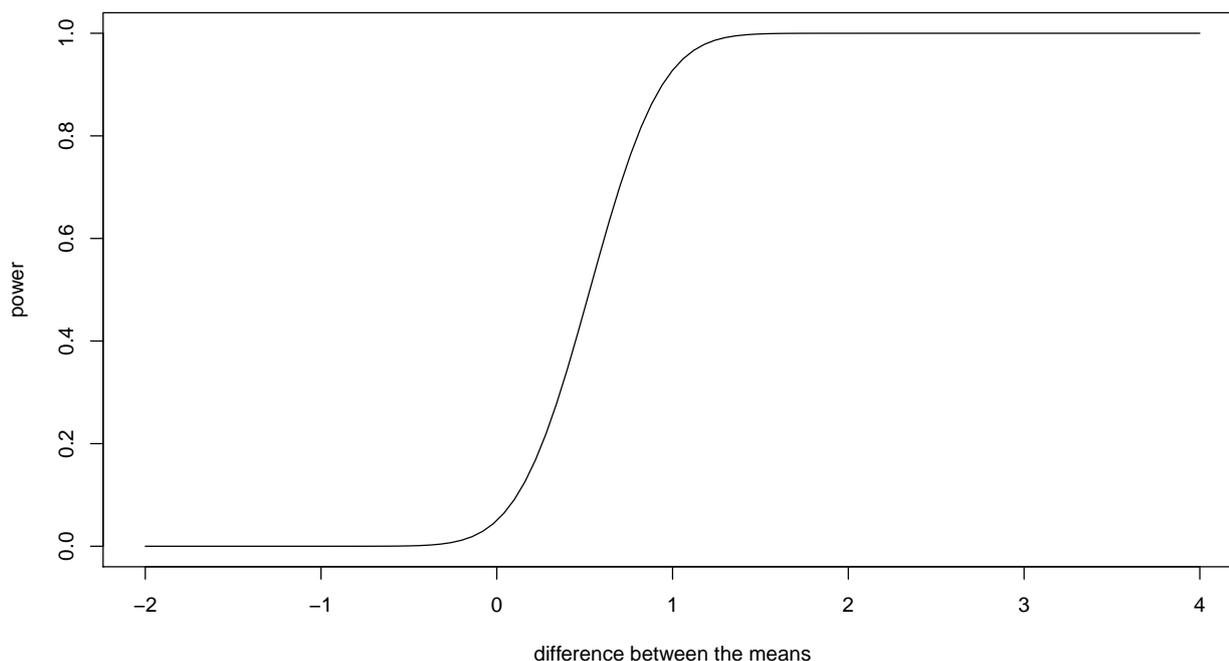
Figure 1: Power of a one-sided two-sample t-test as function of the true difference between the means, $\mu_1 - \mu_2$, assuming that $\sigma = 1$ and that the sample size $n = 20$.

```
    X2 <- rnorm(20,mu2,1)
    t.statistic <- t.test(X1, X2, alternative="greater",var.equal=T)$statistic
    if(abs(t.statistic)>qt(0.95,df=38))       # df = 2n-2
      n.rejections <- n.rejections+1
  }
  n.rejections/m
}
```

2. The power of the test if the true difference between the means is zero, is equal to the significance level α.

```
> power.t.test(n=20,delta=0,sd=1,type="two.sample",
alternative="one.sided")

    Two-sample t test power calculation

              n = 20
          delta = 0
             sd = 1
      sig.level = 0.05
          power = 0.05
    alternative = one.sided

  NOTE: n is number in *each* group
```

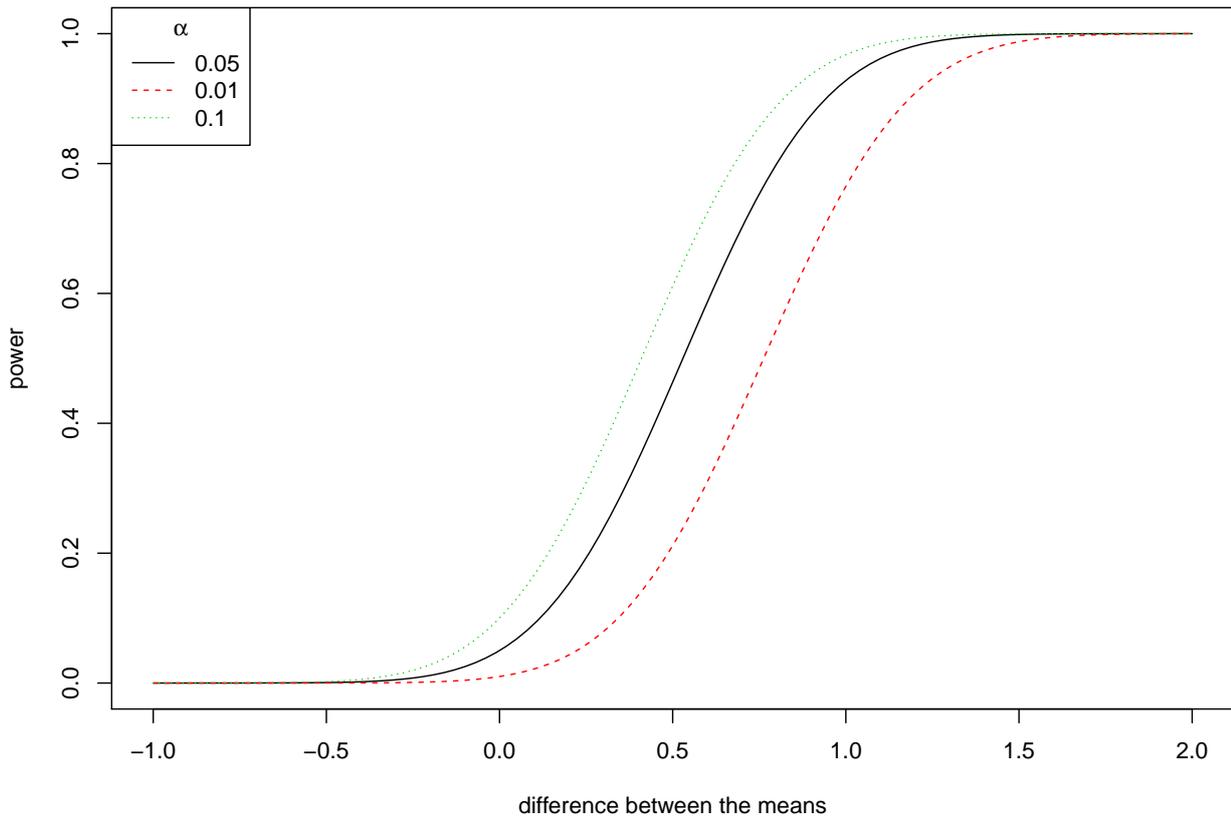3. curve(power.t.test(n=20,delta=x,sd=1,type="two.sample",

Figure 2: Power of a one-sided two-sample t-test as function of the true difference between the means, $\mu_1 - \mu_2$, assuming that $\sigma = 1$ and that the sample size $n = 20$, for different values of α.

```
alternative="one.sided", sig.level=0.01)$power, from=-1, to=2)
curve(power.t.test(n=20,delta=x,sd=1,type="two.sample",
alternative="one.sided", sig.level=0.01)$power, from=-1, to=2,
add=T, lty=2,col=2)
curve(power.t.test(n=20,delta=x,sd=1,type="two.sample",
alternative="one.sided", sig.level=0.1)$power, from=-1, to=2,
add=T, lty=3,col=3)
```

Recall that the power of a given test is defined as the probability of rejecting the null hypothesis H0, given that the alternative hypothesis H1 is correct (see table below). For a given true difference between the means, $\mu_1 - \mu_2$, increasing the significance level α will increase the power of the test. Thus, the higher α, the larger the probability to reject the null hypothesis H_0 given that the alternative hypothesis H_1 is correct (power), but also the larger the probability to reject H_0 given that H_0 is true (type I error). A smaller α will decrease the probability of falsely rejecting H_0, but at the same increase the probability of falsely accepting H_0 (imagine shifting the middle line in the table to the left or the right).

		H_0	
		accept	reject
H_0	True		α type I error
	False	β type II error	power